DOCUMENT RESUME

ED 480 902 TM 035 255

AUTHOR	Matsumura, Lindsay Clare
TITLE	Teachers' Assignments and Student Work: Opening a Window on Classroom Practice. CSE Report.
INSTITUTION	National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.; California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	Stuart Foundation, San Francisco, CA.; Annenberg Foundation, St. Davids, PA.; Los Angeles Unified School District, CA.; Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	CSE-R-602
PUB DATE	2003-07-00
NOTE	50p.
CONTRACT	R305B960002;PR/No.98-85;PR/No.014143;PR/No.0100786;R305B960002
PUB TYPE	Information Analyses (070)
EDRS PRICE	EDRS Price MF01/PC03 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Assignments; *Educational Practices; Elementary Secondary Education; Interrater Reliability; *Teachers; *Teaching Methods

ABSTRACT

This report describes 4 years of research by the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) on developing indicators of classroom practice that have the potential to be used in large-scale settings and that draw attention to important aspects of standards-based learning and instruction. CRESST's method was based on the collection of teachers' assignments with student work. The assignments were then rated and results were summarized to create indicators of classroom practice. Results to date indicate an acceptable level of interrater reliability across study years. It likely would be necessary to collect as many as three or four assignments from teacher to obtain a stable estimate of quality. Additionally, this method was reliable when teachers created their own assignments, but not when teachers submitted assignments created by outside sources. The quality of classroom assignments was associated with the quality of observed instruction, as well as the quality of students' written work. Students who were exposed to teachers who created more cognitively challenging assignments and who had cleared grading criteria also made greater gains on the Stanford Test of Achievement, Ninth Edition (Stanford 9). The quality of teachers' assignments submitted at each of the study years, however, tended to be of basic quality only. Teachers' reactions to the data collection and implications for the use of this method in collaborative professional development sessions also are discussed. (Contains 5 figures, 1 table, and 73 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

CRESST

National Center for Research on Evaluation, Standards, and Student Testing

ED 480 902

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

K. Hurst

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Teachers' Assignments and Student Work: Opening a Window on Classroom Practice

CSE Report 602

Lindsay Clare Matsumura
CRESST/University of California, Los Angeles

TM035255

UCLA Center for the Study of Evaluation

In Collaboration With:

• The RAND Corporation

BEST COPY AVAILABLE

Teachers' Assignments and Student Work:
Opening a Window on Classroom Practice

CSE Report 602

Lindsay Clare Matsumura
CRESST/University of California, Los Angeles

July 2003

Center for the Study of Evaluation
National Center for Research on Evaluation,
Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project School and Classroom Practices in LAAMP Schools
Lindsay Clare Matsumura, Project Director, CRESST/University of California, Los Angeles

Copyright © 2003 The Regents of the University of California

The work reported herein was supported by the Stuart Foundation, PR/Award Number 98-85, with additional funding provided by the Los Angeles Annenberg Metropolitan Project, PR/Award Number 014143; by the Los Angeles Unified School District, PR/Award Number 0100786: Amd. No. 1; and by the Educational Research and Development Centers Program, PR/Award Number R305B960002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this publication do not reflect the positions or policies of the Stuart Foundation, the Los Angeles Annenberg Metropolitan Project, the Los Angeles Unified School District, or the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U.S. Department of Education.

TEACHERS' ASSIGNMENTS AND STUDENT WORK: OPENING A WINDOW ON CLASSROOM PRACTICE

Lindsay Clare Matsumura

CRESST/ University of California, Los Angeles

Jenny Pascal

CRESST/ University of California, Los Angeles

Abstract

In this report, four years of CRESST's research is described developing indicators of classroom practice that have the potential to be used in large-scale settings and that draw attention to important aspects of standards-based learning and instruction. CRESST's method was based on the collection of teachers' assignments with student work. The assignments then were rated and results were summarized to create indicators of classroom practice. Results to date indicated an acceptable level of inter-rater reliability across study years. It likely would be necessary to collect as many as three or four assignments from teachers to obtain a stable estimate of quality. Additionally, this method was reliable when teachers created their own assignments, but not when teachers submitted assignments created by outside sources. The quality of classroom assignments was associated with the quality of observed instruction, as well as the quality of students' written work. Students who were exposed to teachers who created more cognitively challenging assignments and who had clearer grading criteria also made greater gains on the Stanford Test of Achievement, 9th Edition (Stanford 9). The quality of teachers' assignments submitted at each of the study years, however, tended to be of basic quality only. Teachers' reactions to the data collection and implications for the use of this method in collaborative professional development sessions also are discussed.

Introduction

The quality of public education in the United States has come under extreme scrutiny. In addition to concerns raised regarding persistently low academic achievement for low-income and minority students, the state of education generally has been criticized based on large-scale assessment studies (see for example, Gentile, 1992) and studies comparing education in the United States to other industrialized countries (Stevenson & Stigler, 1992). In response to these concerns, the number of school reform activities taking place in public schools has steeply increased. These activities have ranged from the development of new curricula and assessment methods to new professional development settings for teachers.

Chief among these reform activities over the past decade has been the development of content and performance standards for learning and instruction, currently adopted in 49 states (Rothman, Slattery, Vranek, & Resnick, 2002). The standards-based approach to education is based on the idea that nearly all children can master a challenging curriculum and should have the opportunity to do so (Smith & O'Day, 1990, as cited in Rothman et al., 2002). The foundation of this approach is the idea that student achievement will improve as a result of higher quality instruction supported by curricula, assessment strategies, and professional development activities that are aligned with content and performance standards (Briars & Resnick, 2000; Rothman et al.). Many barriers exist to the implementation of standards into everyday classroom activities, however, which severely impede the potential of standards for improving student learning. These barriers to improving instructional quality include the fact that some content standards are framed in relatively general terms and so provide insufficient information to teachers with regard to the recommended content and structure of their learning activities. Also, few professional development opportunities and tools exist that help teachers translate standards into classroom practice (Briars & Resnicku; Rothman, et al.).

In addition to these barriers, the nature of the assessments used in accountability systems also could be considered a barrier to the implementation of standards. Research indicates, for example, that the assessments used to measure student outcomes are not always well aligned with standards (Rothman et al., 2002). Additionally, the success of content standards (and other reform policies) generally has been assessed in one way—through student outcome scores on standardized

tests of achievement. Less emphasis has been placed on assessing the quality of instruction and the ways in which the classroom learning environment may (or may not) transform over time. This is a critical problem given that instructional quality is the most important school factor that influences student learning (Darling-Hammond, 2000; Tharp & Gallimore, 1988). The result is that little information has been available to policymakers, school officials, and teachers regarding the implementation and effect of content standards and other reform activities on students' opportunity to learn in classrooms.

Why Indicators of Classroom Practice Are Needed

One reason that the quality of instruction has remained a "black box" in many accountability systems and large-scale evaluation designs is because few assessment tools exist that directly measure the quality of classroom practice on a broad scale. Teacher surveys frequently have been used to indirectly assess the quality of students' learning environments, though this method has limitations as far accurately describing the interactions between teachers and students, as well as teachers' translations of reform policies (including content standards) into everyday classroom practice (Mayer, 1999; Spillane & Zueli, 1999). Likewise, analyses of student work have provided some information about student performance, but have not drawn attention to the opportunities students have in the classroom to produce high-quality work. Classroom observations have been the most direct way to measure instructional quality, but these can be time consuming and expensive to conduct.

New indicators that help schools, districts, and states monitor and support efforts to improve the quality of instruction are clearly needed. These indicators must also provide information to schools and districts about their interim progress toward reform goals. This is especially important given the fact that numerous studies have shown that even when teachers "buy-in" to a change in instructional practice, the classroom implementation of such a practice does not always reflect a reform program's goals. This is true for subjects as diverse as teaching mathematics from a more conceptual perspective, to implementing the process approach to writing instruction (Applebee, 1984; Cohen & Ball, 1994; Matsumura, Patthey-Chavez, Valdés, & Garnier, in press; Spillane & Zeuli, 1999). More specifically, research indicates that this may be true as well with regard to the implementation of content standards for learning and instruction (Briars & Resnick, 2000).

In this report, findings from four years of CRESST research are described regarding the development of indicators of classroom practice that can be used in large-scale data collections and that draw attention to important aspects of learning and instruction. The unique CRESST methodology includes collecting a sample of teachers' assignments and associated student work and then applying a standardized rubric for describing the quality of the assignments. The results of this scoring process are then summarized to create indicators of classroom practice.

The first section of this report comprises research looking at instructional quality generally, and assignment quality more specifically. Although not intended to be restricted to a single subject area, the work to date has centered on the collection of language arts assignments and the standards for student learning set out in the *Reading/Language Arts Framework for California Public Schools* (California Department of Education, 1999). Findings then are described regarding the psychometric quality of the CRESST classroom assignment rubric, the feasibility of collecting assignments from teachers, and the relation of assignment quality to improved student learning. Also explored are issues relating to how this measure could be used to support on-site collaborative professional development for teachers.

Features of Standards-Based Quality Instruction

While the difficulty of reforming schools has been well documented (see for example, Fullan, 2000, and Tyack & Cuban, 1995), a growing body of evidence indicates that quality instruction positively influences student learning and may be the most important *school* factor influencing student achievement (see for example, Darling-Hammond, 2000; Newmann, Marks, & Gamoran, 1996; Newmann, Bryk, & Nagaoka, 2001; Saunders & Goldenberg, 1999; Tharp, 1982; Tharp & Gallimore, 1988). For example, results from the Tennessee Value-Added Assessment System indicated that teacher effectiveness was the single largest factor that influenced gains in student achievement, an influence that was much larger than poverty or per pupil expenditures. The features of effective instruction include providing students with meaningful and cognitively challenging learning activities as well as opportunities to display their understanding through extended responses. Effective teachers also hold clear goals for student learning and provide students with substantive and specific feedback on their learning. These features of effective instruction are described in more detail in the following sections.

Cognitively challenging and meaningful instruction. Research focused on identifying the elements of quality instruction has found that effective teachers balance direct teaching of skills and concepts with opportunities for students to develop higher level cognitive skills (Newmann, Marks, & Gamoran, 1996; Porter & Brophy, 1988; Resnick, 1999; Slavin & Madden, 1989). Higher order (or independent, symbolic) thinking skills are exemplified by students who construct or manipulate information and ideas. This includes synthesizing, interpreting, evaluating, comparing, etc., or "arriving at conclusions that produce new meanings or understandings for them" (Newmann, Marks, & Gamoran, p. 289). These skills can be developed, for example, in activities in which students apply information to new contexts (e.g., solve a problem), construct arguments, or consider alternative perspectives (Newmann, Marks, & Gamoran; Spillane & Zeuli, 1999). Lower-order thinking, in contrast, occurs when students recite basic factual information or employ a standard, preordained rule or format (Newmann & Wehlage, 1993).

In order to develop students' thinking skills, effective teachers were found to provide students with opportunities to engage with core academic content material with sufficient complexity or "grist" to deeply engage students (Beck & McKeown, in press; Resnick, 1999). As described by Perkins and Blythe (1994), while any topic can be "taught for understanding" by a good teacher, some topics engage students more thoughtfully in subject matter understanding. Specifically, these topics were central to a discipline, accessible to students, and connected to diverse topics within and outside a discipline. Effective teachers also were found to sustain longer term and deeper examination of topics, rather than superficial coverage (Brophy, 1992; Onosko, 1992; Prawat, 1992).

Effective teachers also provide students with opportunities to display understanding through extended responses. For example, researchers investigating classroom conversations found that effective teachers actively elicited more extended student contributions (Beck & McKeown, in press; Goldenberg, 1993; Palinscar, 1986; Tharp & Gallimore, 1988) and allowed sufficient "wait-time" for students to express their ideas (Onosko, 1992). Effective teachers also built on student contributions, asked fewer "known-answer" questions, and elicited basis for statements or positions (Beck & McKeown, in press; Goldenberg, 1992-1993; Henningsen & Stein, 1997; Palinscar, 1986; Tharp & Gallimore, 1988). Teachers used these conversations to make higher order thinking transparent to students by expressing their own thinking processes or problem-solving strategies out loud.

Finally, challenging and meaningful instruction has been described in terms of teachers using real world and students' own experiences to frame lessons and discussions. For example, drawing in part on Vygotsky's (1978) developmental theories, some researchers have characterized effective instruction as drawing on, or "activating" students' background knowledge of a subject (Goldenberg, 1992-1993; Palinscar, 1986; Patthey-Chavez & Clare, 1996; Tharp & Gallimore, 1988). Other researchers have asserted that effective instruction engages students in activities that have meaning beyond the school context (Newmann & Wehlage, 1993), and more closely resembles the kinds of real world problems encountered within a discipline (Henningsen & Stein, 1997).

Clear goals for student learning. Differences in the goals teachers hold for students also help explain differences in teachers' levels of effectiveness (Porter & Brophy, 1988). Specifically, effective teachers have been found to hold clearly articulated goals (Slavin & Madden, 1989) that emphasized conceptual understanding (Brophy, 1992; Onosoko, 1992). Effective teachers also were able to provide more elaborate and detailed descriptions of their instructional goals than less effective teachers (Onosoko). Less effective teachers, in contrast, were found to hold goals that emphasized "teacher transmission" of ideas rather than student application of higher order thinking skills (Onosko). Less effective teachers also tended to focus more on the means by which instruction would take place (an activity) rather than the ends, resulting in an emphasis on student participation (e.g., completion of tasks) rather than on mastery of concepts (Clark & Yinger, 1977; Duffy, 1981; Onosoko).

Substantive and specific feedback. In addition to creating challenging learning environments and holding clear instructional goals focused on student learning, effective teachers were found to use high-quality assessment criteria that were aligned with their goals for student learning and the instructional activity (Black & Wiliam, 1998; Perkins & Blythe, 1994). Effective teachers communicated to students what was expected of them and why, assessed student needs, and adapted their instruction to meet those needs (Black & Wiliam; Perkins & Blythe; Porter & Brophy, 1988). Effective teachers also may communicate to students the criteria by which their performance would be assessed in advance of their completing a task, so that these criteria could be used by students to improve their performance (Goodrich, 1996; Project Zero, 2000; Resnick, 1995).

Classroom Assignment Quality Aligned With Standards

The CRESST method for looking at assignment quality was based on the research described here that focused on instructional effectiveness and reflected a standards-based approach to instruction (California Department of Education, 1998; 1999; Danielson, 1996). We drew as well from the work of other researchers who have examined assignment quality (Newmann, Lopez, & Bryk, 1998; Peterson, 2001; Rademacher, Cowart, Sparks, & Chism, 1997).

Most research to date has focused on the level of cognitive challenge and "authenticity" of English language arts assignments, in addition to the clarity of the assignment directions. Specifically, researchers have defined high-quality assignments as those that provide students with an opportunity to construct knowledge (i.e., develop higher-order thinking skills), draw on a prior knowledge base, develop in-depth understanding, and write extended responses (Newmann, Lopez, & Bryk, 1998; Peterson, 2001). Researchers also have described high-quality assignments as those that provide students with an opportunity to compose for an authentic audience (i.e., to convey information the reader does not know) and have value beyond the immediate school context (Newmann, Lopez, & Bryk; Peterson). Other features of high-quality assignments include clear, detailed, and complete guidelines for students that break the task down into its constituent parts and allow students some degree of choice over features of their work (Peterson; Rademacher, Cowart, Sparks, & Chism, 1997).

CRESST focused as well on the level of cognitive challenge posed by assignment tasks. In addition to cognitive challenge, however, CRESST considered teachers' instructional goals and the grading criteria they used to assess students' work. Specifically, CRESST's framework presupposed that teachers who created high-quality assignments would have clear instructional goals focused on student learning and meaningful content, and that these goals would be carried out in the implementation of the assignment. CRESST's framework also presupposed that high-quality assignments would be cognitively rigorous, based on substantive content material, would require students to produce elaborated responses, and would be aligned with standards. High-quality assignments also would have clearly articulated assessment criteria that would provide substantial information to students regarding what they needed to do to successfully complete the task. These criteria also would be tightly aligned with learning goals so teachers could better

monitor students' progress toward the attainment of core academic concepts and skills. The specific dimensions used to describe assignment quality are presented below (each dimension was rated on a 4-point scale, 1 = *poor* to 4 = *excellent*).

Cognitive challenge. This dimension describes the level of thinking required of students to complete the task. Specifically, this scale focuses on the degree to which students have the opportunity to apply higher order thinking skills (i.e., construct or transform knowledge) and support their answers using evidence from a text. For example, this might mean that seventh-grade students identify and analyze themes across works, analyze characterization, or contrast points of view with a focus on deeper level meanings of texts (California Department of Education, 1999). This dimension also considers the degree to which students engage with core academic content (e.g., read grade-appropriate books as suggested in the *Recommended Readings in Literature, Kindergarten through Grade Eight*[1]). Finally, this dimension considers whether students are required to produce extended responses (i.e., that students at fourth grade and above write multi-paragraph essays) (California Department of Education, 1999). For example, an assignment for which seventh-grade students were asked to write a five-paragraph essay comparing themes across grade-appropriate books would likely receive a high score for cognitive challenge. An assignment given a low score on this dimension, in contrast, might require students to recall very basic, surface-level information (e.g., one- to two-sentence responses to questions such as, "What color was the car?") or write on a topic requiring no academic content knowledge (e.g., a fan letter to a movie star).

Clarity of the learning goals focused on student learning. This dimension describes how clearly a teacher articulates the specific skills, concepts, or content knowledge students are to gain from completing the assignment. The purpose of this dimension is to describe the degree to which an assignment could be considered a purposeful, goal-driven activity focused on student learning, rather than "activity for activity's sake." An assignment given a high score on this dimension would have goals that were very clear, detailed, and specific as to what students were to learn from completing the assignment. It would also be possible to assess whether or not students had achieved these goals. For example, the following set of goals for a third-grade assignment received a high score: "We expected the students to continue developing paragraphs that develop a central idea. It was expected that they stay on

---

[1] California Department of Education (1996). *Recommended Readings in Literature, Kindergarten through Grade Eight*. Sacramento: California Department of Education.

the topic, give details, and show awareness of audience." The teacher goals for another third-grade assignment, in contrast, were much less specific: "I wanted [the students] to properly express their ideas and answer the prompt correctly."

Clarity of the grading criteria. The purpose of this dimension is to assess the quality of the grading criteria teachers use to assess student work. How clearly each aspect of the grading criteria is defined is considered in the rating, as well as how much detail is provided for each of the criteria. An assignment given a high score on this dimension would have a grading criteria in which the guidelines for success were clearly detailed and provide a great deal of information to students about what they needed to do to successfully complete the task. Most of the assignments that received a high score on this dimension used writing rubrics with well-differentiated and elaborated score points focused on a number of critical aspects of writing, not just surface-level mechanical features.

For example, one assignment that received a high score on this dimension included a rubric that consisted of three dimensions measuring different aspects of students' written work: writing strategies, writing application, and writing conventions. Each dimension was assessed on a 4-point scale (1 = *beginning* to 4 = *exceeds standards*). Each scale point for each dimension of this rubric gave detailed, elaborated information about what was expected in students' writing. On the other hand, a moderate score was assigned to teachers who provided a list of features upon which student work was graded (e.g., clarity, spelling, grammar, awareness of audience, and examples from story), but did not specify a range of success for each feature. In other words, assignments such as these would have received a higher score for this dimension if teachers had specified how many examples students were expected to include in their writing in order to receive a high, medium, or low grade.

Alignment of goals and task. This dimension focuses on the degree to which a teacher's stated learning goals are reflected in the design of the assignment tasks students are asked to complete. Specifically, this dimension attempts to capture how well the assignment appears to promote the achievement of the teacher's goals for student learning. An assignment given a high score on this dimension would have learning goals and tasks that overlapped completely. For example, a third-grade teacher's goals for one assignment were that students analyze a text (*The Hundred Dresses*, by Eleanor Estes) and connect it to their own lives. Students were required to write extended responses describing which characters in the story they most

identified with and what connections they could make between the story and their own lives. A teacher's goals for an assignment given a low score on this dimension, in contrast, were that students would "connect what they read to their own experience" and learn to "appreciate the ideas of others." Students read *Chicken Sunday* by Patricia Polacco, a story in which three children raise money to buy a hat for their grandmother to thank her for her chicken dinners on Sundays. The actual assignment task, however, required students to write a description of a project for which they needed to raise money by selling decorated eggs. The link to the rich content of the story, which addresses both cultural and intergeneration issues, was superficial and indirect. Additionally, there was no evidence that the task helped promote the students' understanding and appreciation of the ideas of others.

Alignment of goals and grading criteria. This dimension is intended to describe the degree to which a teacher's grading criteria support the learning goals, that is, the degree to which a teacher assesses students on the skills and concepts they are intended to learn through completion of the assignment. Also considered in this rating is whether or not the grading criteria include extraneous dimensions that do not support the learning goals, as well as the appropriateness of the criteria for supporting rigorous, standards-based learning goals. An assignment given a high score on this dimension would have goals and grading criteria that overlapped completely. An assignment given a low score on this dimension, in contrast, would have grading criteria that did not support the learning goals. One assignment that received a high score on this dimension included the teacher's goal that students improve their writing skills by interviewing their peers and writing up the results of their interviews and that the students learn to distinguish good interview questions from bad ones. The grading criteria the teacher used to measure the attainment of these goals focused on the extent to which the students' writing contained "factual information [gathered] using the question-answer-form," and the extent to which "readers [got] to know the person interviewed through the questions the student asked."

The teacher's goals for an assignment that received a low score on this dimension, in contrast, were to improve students' reading comprehension, and to teach them to answer comprehension questions fully and in detail. In describing her grading criteria for the assignment, however, the teacher wrote only "Each question is worth 20 points. Partial answers 10 points. Doesn't apply 0 points." The teacher's grading criteria did not reference the skills that were directly connected to her goal

that the students develop reading comprehension skills (e.g., the ability to provide a complete plot summary, identify a theme, make predications based on previous events in the story, etc.). Also, her criteria did not provide students with information that would help them complete the assignment task successfully.

Overall quality. This dimension is intended to provide a holistic rating of the quality of an assignment based on its level of cognitive challenge, clarity of the teacher's learning goals, clarity of the grading criteria, alignment of the learning goals and task, and alignment of the learning goals and the grading criteria.

CRESST also is continuing to develop other dimensions of quality for looking at assignments and student work. Based on the work of the Educational Testing Service and NAEP, one of these new dimensions may focus on the degree to which teachers provide students with clear and elaborated assignment directions (described in Peterson, 2001). Also, beyond focusing on the clarity of teachers' goals, it may be important to look at whether teachers are basing those goals on standards. Just looking at whether teachers cite standards, however, likely would not be enough. Research indicates that there can be a tendency to focus on less challenging standards when designing assessments (and this could be true when designing assignments as well) (Rothman et al., 2002). For this reason, it likely would be important to focus on the degree to which teachers' goals emphasize students' attainment of complex thinking skills as well.

Finally, in addition to looking at how clear and informative teachers' grading criteria are to students, it may also be important to look at the degree to which teachers assess students on standards-based skills and complex thinking skills. Research indicates that teachers frequently emphasize the mechanical features of students' writing in their feedback to them and do not provide students with substantive feedback on the content of their work. When students are provided with more substantive feedback, however, their work shows improvement across drafts (Matsumura et al., in press). Based on this research, and other research focused on the quality of classroom assessments and student learning (cited in Black & Wiliam, 1998), it likely would be important to consider the content of teachers' grading criteria in some way in addition to clarity. These dimensions are still under development, however, and have not yet been validated.

Findings and Lessons Learned From CRESST's Research and Development Work

During the first two years of the study, CRESST focused on four elementary and four middle schools that were part of evaluation of the Los Angeles Annenberg Metropolitan Program. These schools served primarily poor and minority students, the majority of whom were English language learners. Four to six language arts assignments were collected from third- and seventh-grade teachers (n = 24). These assignments included "typical" writing and reading comprehension work, as well as assignments identified by teachers as being "challenging" for the students in their class (Aschbacher, 1999; Clare, 2000; Clare & Ashbacher, 2001).

At year three, third grade was again targeted for study, but the research was scaled up to include elementary schools that served primarily middle-class white and Asian students (n = 29 teachers). The purpose of this was to see if CRESST's method of looking at assignments and student work would continue to serve as an effective indicator of instructional practice across a wide range of classroom learning environments. Only two assignments were collected from teachers that year in order to investigate whether the few assignments could yield a stable estimate of quality (Clare, Valdés, Pascal, & Steinberg, 2001; Matsumura et al., in press). Teachers at all three study years were paid a stipend of $100 for their participation.

At the fourth year of the study (the 2000–2001 academic year), CRESST collaborated with the Los Angles Unified School District (LAUSD) to pilot the assignment rubric as part of its new accountability system. The LAUSD recently had divided into 11 "local districts," each with its own superintendent. This system was developed to monitor the progress of the new local district superintendents. Building on early systems of accountability, LAUSD's new system also sought to measure both direct outcomes of student performance and the school processes expected to increase student performance (Cantrell, Lyon, Valdés, White, Recio, & Matsumura, 2001). Two of the four indicators that made up this new accountability system were intended to measure student performance. The remaining two indicators (including the CRESST classroom assignment measure) were intended to measure important school and classroom processes that potentially could influence student achievement.

For the LAUSD classroom assignment pilot, 181 fourth-, seventh- and tenth-grade teachers in 35 schools were randomly selected from the 11 local districts to participate. Teachers were given three weeks to submit three assignments (one

typical writing assignment and two typical reading comprehension assignments), though the deadline for submissions was later extended. Teachers also were given a copy of the assignment rubric and were asked to complete a short survey describing their reactions to the assignment measure and data collection. Of the teachers who were sampled, 50 returned completed assignment materials (a 28% participation rate). Teachers were given $75 for classroom materials for their participation.

At each study year, teachers were asked to complete a one-page cover sheet describing their learning goals and assessment criteria (see Appendix for the cover sheets used in the LAUSD data collection). Teachers also submitted four samples of student work for each assignment—two of which they considered to be of high quality, and two of which they considered to be of medium quality. Teachers also were observed in their classrooms twice in years 2 and 3 of the study, and once for the LAUSD pilot study. Students' work was assessed at each study year using rubrics that were developed by CRESST and the United Teachers of Los Angeles and that measured the content, organization, and mechanics of students' writing (Higuchi, 1996).

Reliability and Validity of the Classroom Assignment Ratings

To explore the psychometric quality of the classroom assignment rubric, agreement between raters was investigated at each of the study years. The quality of classroom assignments and observed instruction also were compared in order to look for evidence of the construct validity of the assignment ratings.

Inter-rater reliability. Results across all four years of the study indicated an acceptable level of agreement between raters overall. Cohen's kappa coefficients were calculated at each year to investigate whether the pattern of agreement observed was greater than would be expected if the raters had randomly assigned scores. Kappa coefficients for each dimension for each assignment were significant at the $p < 0.01$ or higher and of a moderate magnitude at each year of the study. Alpha coefficients also indicated an acceptable level of internal consistency for each dimension at each study year, and the percentage of agreement between at least two CRESST raters was greater than 80% for each dimension.

When the pool of raters was expanded for the LAUSD pilot study, results indicated that, while overall agreement was acceptable, the percent of exact scale agreement between *individual* rating pairs ranged considerably. For example, the

correlation between the novice raters and the expert raters on ratings of cognitive challenge ranged from $r = .92$ to $r = .41$ for the elementary school assignments, and $r = .82$ to $r = .30$ for the secondary assignments. Not surprisingly, the two expert raters who had been part of CRESST's previous research had the highest level of agreement. The novice raters with the highest level of agreement with the CRESST raters had some experience as classroom teachers combined with some background in educational evaluation. The raters with only classroom teaching experience, in contrast, had the lowest level of agreement with the expert raters.

Assignment ratings and observed instruction. Classroom assignment ratings also were compared with ratings of observed instruction to investigate the degree to which the classroom assignment ratings yielded meaningful and appropriate information about students' learning environments that was commensurate with other measures of quality practice. Results indicated that the classroom assignment ratings were associated with the quality of observed instruction across the CRESST study years, especially with regard to the level of cognitive challenge of the observed lessons and assignment tasks (Clare & Aschbacher, 2001; Clare et al., 2001).

For the LAUSD pilot study, observations were conducted by district research staff (using an abbreviated form of the CRESST observation protocol). Contrary to what we had found at previous years, results indicated that the quality of the assignment ratings was not associated with the quality of observed instruction for those few teachers who were observed and who submitted assignments ($n = 16$). These last analyses, however, were limited by the small sample size.

In summary, rater reliability was acceptable across study years, though future efforts to train inexperienced raters should include a more substantive focus on strategies for applying rubric scores (i.e., should stress the fundamental uses and limitations of rubrics). Additionally, it might be necessary to screen out raters with low levels of interrater agreement. The classroom assignment ratings also were significantly associated with the quality of observed instruction across most of the study years and across a range of learning environments, providing evidence for the construct validity of this method.

The Feasibility of Collecting Assignments

In addition to investigating the reliability and validity of our method, the potential feasibility of collecting assignments on a large-scale basis also was examined. Teachers generally reported that it took at least an hour per assignment to

complete the cover sheet and choose and Xerox student work, and some reported that it took even longer. Teacher time and burden was an extremely important issue to consider, given that these (and many other urban) schools were involved in multiple reform projects, each of which often required its own set of evaluation activities. A significant focus of CRESST's research, therefore, was on identifying the minimum number of assignments needed to obtain a stable estimate of the quality of classroom practice.

Estimating the number of assignments to collect from teachers. Generalizability analysis techniques that estimate the relative magnitude of different components of error variation were used at each year of the study to investigate the number of assignments and raters needed to yield a stable estimate of quality (Shavelson & Webb, 1991). Results indicated that the collection of four assignments from teachers rated by three raters yielded a G-coefficient of .91 for elementary school and .87 for middle school (.80 and above is considered to be good) (Clare & Aschbacher, 2001). The following year of the study, two assignments were collected from teachers resulting in a G-coefficient of only .64, an unacceptable level of stability.

For the LAUSD pilot study, teachers were asked to submit three assignments. The elementary- and secondary-level assignments then were rated independently by three novice raters and two experienced CRESST raters. A larger pool of raters was used in order to investigate the rater reliability of individuals with varying backgrounds (these results were described earlier).

Results indicated that the collection of three assignments from teachers yielded a stable estimate of quality at the secondary level (G = .88) and that most of the variation in assignment quality was *between* teachers. At the elementary school level, in contrast, this design yielded a G-coefficient of only .46, and most of the variation in assignment quality was *within* teachers. In other words, individual teachers at the elementary level of schooling tended to submit assignments of varying quality, whereas the secondary teachers tended to submit assignments that were more similar in quality. These results may have been partially due to the fact that many of the elementary school teachers submitted commercially produced assignments as well as assignments they created on their own. Specifically, 27% of the writing assignments and 59% of the reading comprehension assignments from the elementary school teachers were generated from outside sources. This contrasted

with secondary teachers (and the teachers in past years of data collection) who submitted almost all teacher-created assignments.

Teachers' perspective on the data collection. To further investigate the feasibility of the data collection, teachers in the LAUSD study were asked to provide feedback on the data collection by completing a very short (one-page) survey. Approximately 80% of the teachers who completed the survey (n = 38), agreed that the rubric and data collection process supported reflection on the quality of their assignments, the nature of their learning goals, and the quality of student work in their classroom. In the words of one teacher:

> It causes one to really pause and reflect on the tasks one assigns to students. Are they complex? Focused? Are grading criteria explicit? Clear? Are learning goals aligned? It's a motivator for improvement should one be really honest with oneself. My biggest criticism: The coversheets alone were very time-consuming [to complete]. You need to realize how much teachers have to do! I did most of this on my weekends, without pay.

While teachers were generally positive about the data collection (at least the 28% who participated in the LAUSD pilot study), more than half of the teachers reported that they needed more than the amount of time allotted by the district to complete the assignment materials. As one teacher wrote,

> Give us more time! How about initiating the process at the beginning of the school year instead of a mere three weeks before you expect a completed package sent back! Give us that professional courtesy please!

A few teachers also suggested that the data collection not be done around a holiday period (n = 4).

In summary, it likely would be necessary to collect as many as three or four assignments from teachers to obtain a stable estimate of quality. Future research also is needed to determine whether CRESST's method of collecting and scoring assignments produces a stable and valid indicator of classroom practice when teachers submit commercially produced assignments instead of assignments they create on their own. While teachers (who participated in the LAUSD study) were generally positive about the assignment rubric, they also reported that they needed a significant amount of time to produce assignments and collect student work.

Improving Instruction and Student Learning

Variation in assignment quality was examined at each study year in order to learn more about the diversity of classroom learning environments. The relationship between assignment quality and student work, and the influence of assignment quality on students' achievement test scores, also was examined. The purpose of this was to further investigate the validity of the assignment ratings by examining whether facets of instruction were being measured that were germane to student learning.

The quality of classroom assignments. At each year of the study, the assignments that were collected tended to be of basic quality (i.e., were scored a 2 on a 4-point scale for overall quality). Specifically, results indicated that teachers often did not provide students with the opportunity to apply higher order thinking skills, or engage with substantive content material. Teachers also tended to have nonspecific goals for student learning and grading criteria that provided little information to students regarding what they would need to do to successfully complete the task.

At the elementary school level, reading comprehension assignments typically required students to write short responses to simple, basic recall comprehension questions. For example, one fourth-grade teacher had the students complete a worksheet summarizing the beginning, middle, and end of *Tales of a Fourth Grade Nothing*, by Judy Blume, and answer the following comprehension questions: 1) How does the story compare with real life? 2) Who is the main character? and 3) What similar experiences have you had with your friends? Students wrote one to two sentences for each question. The teacher's goals for this assignment were:

> The students needed to comprehend the story and pull details about the story. They needed to compare their experiences to the characters.

And her grading criteria were:

> My criteria is whether they answered the questions correctly (followed directions) and were they able to include details from the story.

This assignment was scored a 2 for overall quality, because most of the questions required students to summarize straightforward information from the

book. Additionally, though two of the questions asked students to make comparisons between their own lives and the story, students were not expected to elaborate on their responses. For this assignment to receive a higher score on this dimension, the questions would have had to require students to think more deeply about the story, as well as write more extended responses. The teacher's goals for the assignment also were considered to be only of moderate quality because they were primarily stated as activities and did not identify the specific aspects of the story that the teacher believed were important to comprehend. For example, the teacher did not clarify what she wanted the students to learn as a result of comparing their experiences to those of the characters. The teacher's grading criteria similarly were broadly stated. Finally, having students write such short responses was not aligned with the teacher's goal and grading criteria that students include details from the story in their responses.

Assignments tended to be of a similar quality at the middle school level. For example, one typical writing assignment given by a seventh-grade teacher required students to engage in the steps of the writing process and produce a five-paragraph essay on their dreams for the future. The teacher's goals for this assignment were:

> To teach students step by step how to write a five-paragraph essay and to demonstrate the creativity and fun in essay writing.

Her grading criteria for this assignment were:

> Content is explained well. Writers focus on what needs to be talked about. Writing process is completely done.

This writing assignment was scored a 2 for overall quality. Though learning to write a five-paragraph essay is a grade-appropriate task for seventh grade, this assignment would have received a higher score for cognitive challenge if students had been required to draw on substantive content material (e.g., compare themes or characters from books) when writing their essays, or had been required to use their personal experiences to construct a convincing argument. The teacher's goals were quite broad and did not present clear objectives for student learning. Instead, the goals focused on the activity of producing a five-paragraph essay and the application of the basic format of the essay. These goals also were not well aligned with her grading criteria that emphasized that the "content" of the student essays be

"explained well" because her goals did not in fact focus on content. More importantly, however, the teacher did not explain her criteria for determining the degree to which students had explained their content well or focused on "what needs to be talked about." A more specific and elaborated set of grading criteria would have received a higher score and might have helped the students to better understand what was expected of them and how they could have more effectively focused their efforts.

Overall, the quality of the assignments collected from schools serving more-privileged students was statistically of significantly higher quality than the assignments collected in schools serving primarily poor and minority students (Clare et al., 2001; see Table 1). It is important to underscore the fact, however, that there was quite a bit of variation among schools in the quality of the assignments submitted by teachers. In other words, some teachers from schools serving poor students submitted outstanding assignments, while some teachers from schools serving more-privileged students submitted only mediocre assignments.

Classroom assignment quality and standards for student learning. In addition to variation in assignment quality between teachers at the same school who served the same population of students, results indicated that the quality of

Table 1

Quality of Assignments in Classrooms Serving Traditionally Lower- and Higher Achieving Students (N = 29 Teachers)

	Lower achieving (n = 13) *M (SD)*	Higher achieving (n = 16) *M (SD)*	p value
Cognitive challenge of the task	1.64 (.44)	2.23 (.61)	.000
Clarity of learning goals	1.92 (.50)	2.32 (.56)	.007
Clarity of grading criteria	2.37 (1.01)	1.94 (.66)	.07
Alignment of goals and task	1.83 (.49)	2.17 (.48)	.013
Alignment of goals and grading criteria	1.81 (.59)	1.71 (.55)	.52
Overall quality	1.71 (.43)	2.21 (.48)	.000

Note. Items were scored on a 4-point scale (1 = *poor*, 4 = *excellent*).

Reprinted from Clare et al., 2001.

assignments varied quite a bit among teachers who claimed to be adhering to the same set of content standards. Though this was not an explicit focus of the research study, it appears that teachers interpreted and implemented content standards differently in their classrooms. For example, one fourth-grade teacher cited the California standards for reading comprehension (2.0 and 2.1) in describing her assignment (California Department of Education, 1999). These standards state that students should "draw upon a variety of comprehension strategies (e.g., generate and respond to essential questions, make predictions, compare information from different sources)" (California Department of Education, 1999, p. 114). These standards also state that students should "identify structural patterns found in informational text (e.g., compare and contrast, cause and effect, sequential or chronological order, proposition and support) to strengthen comprehension" (Ibid.). The assignment task, however, mostly required students to answer straightforward factual information questions, such as "What does Mrs. F. see as she is crossing the farmyard?" Students also provided a short (one- to two-sentence) response to the open-ended question, "If you were Mrs. F., would you have helped the crow? Why or why not?" This last question was more challenging than the ones that preceded it, but still did not provide students with much of an opportunity to develop the multiple strategies for reading comprehension described in the standards, or identify structural patterns found in informational text (especially since the students read a story).

Other teachers focused on the "writing strategies" portion of the standards, as opposed to content, with uneven results with regard to overall assignment quality. For example, one fourth-grade teacher cited the writing strategies (1.0 and 1.1) standards when describing her assignment (California Department of Education, 1999). These standards state that students should "select a focus, an organizational structure, and a point of view based upon purpose, audience, length, and format requirements" (California Department of Education, 1999, p. 115). Students also are to "create multi-paragraph compositions" that include an introductory paragraph, supporting paragraphs comprised of topic sentences and supporting sentences, and paragraphs that summarize the main points (Ibid.). This teacher had the students read *The House at Pooh Corner*, by A. A. Milne, and write a multi-paragraph story

describing in detail their imagined encounter with one of the story's characters. This assignment was given a relatively high score because students were required to draw on the story character's personality qualities when constructing their stories and provide rich descriptive details. Another teacher who cited the exact same standards, in contrast, received a lower score on her assignment because the students only were required to answer very basic factual questions on a worksheet and write a two-paragraph retelling of the story. It appears that basing assignments on the writing strategies portion of the standards alone does not necessarily result in high-quality assignments that provide students an opportunity to develop complex thinking skills.

Classroom assignments and student learning. Results across all four years of the study indicated that students on the whole benefited from higher quality assignments. Specifically, results indicated at each year of the study that the quality of students' work (notably the quality of the content of students' writing) was significantly associated with the quality of teachers' assignments. In other words, higher quality writing assignments led to higher quality student work and vice versa. Preliminary findings from the LAUSD pilot study also indicated that secondary students showed significant gains on their Stanford 9 reading and language scores when they were exposed to more challenging assignments and assignments with higher quality grading criteria (Matsumura, Garnier, Pascal, & Valdés, 2002). These results are commensurate with the work of other researchers in this area who found that students, even those from very disadvantaged backgrounds, produced higher quality work when they received more cognitively challenging assignments (Newmann, Bryk, & Nagaoka, 2001) and were exposed to higher quality assessments (Black & Wiliam, 1998). Results also indicated a negative relationship between the clarity of teachers' grading criteria and student achievement. These results likely are a result of multicollinearity, however, or a very high degree of association between the predictor variables. On its own, clarity of the teachers' learning goals did not predict student achievement.

In summary, while some teachers submitted outstanding assignments, the majority of the assignments we collected at each study year were only of a basic quality. This was true even when teachers claimed to be adhering to content

standards for instruction. When students were exposed to higher quality assignments, however, they produced higher quality work. Secondary students who were exposed to assignments that were more cognitively challenging and had clearer grading criteria also received higher scores on standardized tests of achievement. These results suggest that improving the quality of teachers' assignments might improve the quality of students' learning environments, though this has not yet been investigated. Further work also may be necessary with regard to revising the dimension that measures the focus of the teachers' goals on student learning (e.g., focusing this more on the content of teachers' goals, etc.).

Finally, these findings raise important issues regarding the implementation of content standards for instruction into classroom practice. Regardless of the quality of content standards (and California's language arts standards are considered to be very high quality overall), there still appears to be quite a bit of room for interpretation of standards and use of standards in practice. For example, California's standards are quite comprehensive and focus on students' development of both higher- and relatively lower level skills. Simply focusing on the parts of the standards that deal with lower level skills, however, would not be in keeping with the spirit and purpose of standards-based education, which is to provide all students with the opportunity to master a challenging curriculum. Additionally, teachers vary quite a bit in how they interpret the meaning of standards. For example, strategies for deepening reading comprehension (e.g., predicting or comparing) can be implemented in more or less challenging ways depending on whether one focuses on surface-level or deeper-level meanings of texts and whether one requires students to provide rationales for their answers. It is possible that reflecting on classroom assignments could help teachers maintain a focus on students' development of complex skills (while attending to the development of lower level skills as well). CRESST's method for looking at classroom assignments also could potentially help teachers critically review the substantive content of their assignments in terms of whether students are being supported to engage with deeper level meanings of texts (in addition to surface-level details), and cite evidence from the text.

Improving the Quality of Teachers' Assignments

In the following sections, some ideas are proposed for how the CRESST method for looking at assignment quality might be used to support teachers' professional development and alignment of their assignments with standards. Factors associated with the effectiveness of collaborative professional development, barriers to implementation, and the protocols currently available to teachers to support work in these settings are described. Specific suggestions then are made for how the CRESST method could be used to help guide teachers' reflection on assignment quality.

Collaborative Professional Development and Student Learning

At the same time that teachers and schools have been held accountable (often publicly) for student success, it is widely acknowledged that traditional models of professional development most available to teachers are inadequate for supporting the improvement of classroom practice (Cohen, McLaughlin, & Talbert, 1993; Lieberman, 1994; Saunders, Goldenberg, & Hamann, 1992). Traditional approaches to professional development expose teachers to new instructional methods and curricula through a limited number of days of in-service workshops that are "unrelated to each other or to the fundamental instructional pedagogical issues teachers face daily" (Fuhrman, 1993, p. 7). Isolated in their classrooms, teachers, for the most part, are then left to interpret programs and new standards for teaching on their own. The end result is that in spite of good intentions, teachers' instructional practices and students' chances for academic success remain essentially unchanged (Cuban, 1990; Tharp & Gallimore, 1988; Tyack & Tobin, 1994). This problem is especially pronounced in urban schools serving poor students because these schools tend to have fewer numbers of well-qualified teachers and greater numbers of students with special learning needs (National Commission on Teaching and America's Future, 1996).

The need for a different approach to teachers' professional development has inspired a number of research efforts focused on the factors that appear to make these settings more effective in terms of improving classroom practice and student achievement (see for example, Newmann & Wehlage, 1993; Stokes, 2001). In brief, effective professional development settings for teachers have been found to be sustained, ongoing, and site-based, and allow teachers to talk with peers about

changes and improvements in their practice (Darling-Hammond & McLaughlin, 1995; McLaughlin & Zarrow, 2001; Powell, Goldenberg, & Cano, 1995; Saunders, Goldenberg & Hamann, 1992). These settings have a member (or members) who serve as coaches (either formally or informally) and provide peers with substantive feedback about their efforts and model excellent teaching strategies (Darling-Hammond & McLaughlin; Powell, Goldenberg, & Cano; Gallimore & Goldenberg, 1992). Finally, activities in these settings focus on an explicit set of goals that guide the group's engagement in joint productive work that focuses on student work and products, as well as the concrete tasks of teaching and assessment (McLaughlin & Marsh, 1990; McLaughlin & Zarrow, 2001; Saunders, Goldenberg & Hamann, 1992).

While the factors associated with effective professional development have been identified, the norms, or culture, of schools in the United States do not readily support teacher collaboration, making it difficult to implement and sustain these types of professional development settings (Hiebert & Stigler, 2000). Many schools, for example, do not have a person on-site who is willing to serve as a coach, or who is skilled enough to model excellent teaching strategies. Teachers also do not always know how to provide each other with substantive feedback or have access to tools that can help them reflect on and assess the quality of their own practice (Matsumura & Steinberg, under review).

To assist the process of implementing these types of settings, a number of different protocols[2] have been developed to support and guide the ways in which teachers interact with each other in collaborative professional development settings (McDonald, 2001). While there has been hardly any systematic research investigating the efficacy of these protocols for supporting improvement in learning and instruction, it seems unlikely that *on their own* these protocols would improve classroom practice because they primarily focus on group dynamics rather than on the concrete tasks of teaching. For example, the Tuning Protocol, developed by Joe McDonald for the Coalition of Essential Schools, and the ATLAS protocol, developed by Eric Buchovecky for the ATLAS community project, focus on helping teachers provide feedback to each other (see Figures 1 and 2). No guidance is given to teachers with regard to the content of their feedback (i.e., what to focus on when contemplating a colleague's lesson or assignments).

---

[2] Some common protocols for specifically looking at student work in collaborative settings include the Collaborative Assessment Conference by Steve Seidel for Harvard's Project Zero; the ATLAS protocol developed by Eric Buchovecky; and the Consultancy, Tuning, and Vertical Slice protocols developed by Joe McDonald at the Coalition of Essential Schools.

The protocol, Standards in Practice, developed by Ruth Mitchell for the Education Trust, focuses on scoring students' work with the intention of drawing attention to standards for student learning (see Figure 3). This protocol is more focused on classroom practice than the Tuning and ATLAS protocols, but still does not draw explicit attention to the opportunity students had to produce high-quality work. For example, the protocol lists "examining assignments to make sure that they are clearly aligned with standards" as an action that could be taken to improve student learning. Explicit guidance for what to look for in assignment quality, however, or how to develop assignments that might better support student attainment of the standards, is not provided.

This does not mean to imply that creating a safe and supportive environment to exchange feedback in a collaborative professional development setting is unimportant. CRESST found, for example, that teachers in these settings cited the support they received from their groups as one of the most important benefits of participation (Matsumura & Steinberg, under review). Similarly, providing guidelines for approaching the scoring of student work also could be very helpful to teachers. Clearly other types of protocols may be needed, however, in addition to existing ones. Specifically, it appears that there is a need for protocols that provide a

I.	**Introduction** [10 minutes]. Facilitator briefly introduces protocol goals, norms and agenda. Participants briefly introduce themselves.
II.	**Teacher Presentation** [20 minutes]. Presenter describes the context for student work (its vision, coaching, scoring rubric, etc.) and presents samples of student work (such as photocopied pieces of written work or video clips or an exhibition).
III.	**Clarifying Questions** [5 minutes maximum]. Facilitator judges if questions more properly belong as warm or cool feedback than as clarifiers.
IV.	**Pause to Reflect on Warm and Cool Feedback** [2-3 minutes maximum]. Participants take note of "warm," supportive feedback and "cool," more distanced comments (generally no more than one of each).
V.	**Warm and Cool Feedback** [15 minutes]. Participants among themselves share responses to the work and its context; teacher-presenter is silent. Facilitator may lend focus by reminding participants of an area of emphasis supplied by the teacher-presenter.
VI.	**Reflection and Response** [15 minutes]. Teacher-presenter reflects on and responds to those comments or questions he or she chooses to. Participants are silent. Facilitator may clarify or lend focus.
VII.	**Debrief** [10 minutes]. Beginning with the teacher-presenter (How did the protocol experience compare with what you expected?") the group discusses any frustrations, misunderstandings, or positive reactions participants have experienced. More general discussion of the tuning protocol may.

Figure 1. "Tuning Protocol" by Joe McDonald for reflecting on student work in collaborative professional development settings (Coalition of Essential Schools, 1996).

When looking for evidence of student thinking:
- Stay focused on the evidence that is present in the work.
- Avoid judging what you see.
- Look openly and broadly; don't let your expectations cloud your vision.
- Look for patterns in the evidence that provide clues to how and what the student was thinking.

When listening to colleagues thinking:
- Listen without judging.
- Tune into different perspectives.
- Use controversy as an opportunity to explore and understand each other's perspectives.
- Focus on understanding where different interpretations come from.
- Make your own thinking clear to others.
- Be patient and persistent.

When reflecting on your own thinking:
- Ask yourself, "Why do I see this student work in this way? What does this tell me about what is important to me?"
- Look for patterns in your own thinking.
- Tune in to the questions that the student work and your colleagues' comments raise for you.
- Compare what your see and what you think about the student work with what you do in the classroom.

When you reflect on the process of looking at student work, ask:
- What did you see in this student's work that was interesting or surprising?
- What did you learn about how this student thinks and learns?
- What about the process helped you see and learn these things?
- What did you learn from listening to your colleagues that was interesting or surprising?
- What new perspectives did your colleagues provide?
- How can you make use of your colleagues' perspectives?
- What questions about teaching and assessment did looking at this students' work raise for you?
- How can you pursue these questions further?
- Are there things you would like to try in your classroom as a result of looking at the students' work?

Figure 2. "ATLAS" protocol by Eric Buchovecky for looking at student work in collaborative professional development settings (Coalition of Essential Schools, 1996).

framework for what to focus on when developing and implementing lesson activities generally, and classroom assignments in specific. In other words, there is a need for tools that help teachers focus on the content and implementation of their assignments, and that draw explicit attention to the opportunity students have to produce high-quality work in classrooms. Additionally, as described earlier, assignments where teachers cited specific content standards were not necessarily cognitively rigorous, nor did these teachers necessarily have high-quality grading criteria for assessing student work. It appears that there also is a need for tools that help teachers create assignments and assessment tools that are more tightly aligned with the meaning and intention of the content standards. Together with the process protocols described earlier, such tools could provide a powerful intervention for improving learning activities.

1. **We all complete the assignment** – Please complete the assignment that the students were asked to do. This is important: If you don't do the assignment yourselves you won't know whether it truly asks for the knowledge and skills you want students to have.
2. **We identify the standards that apply to this assignment** – Identify the standards that apply to this assignment. Take the standards you are using (national, state, local) and find those standards to which this assignment might be directed. In other words, if the students do the assignment, what standards would they be moving toward? (If the answer is "none," then what would be the consequences?)
3. **We generate a rough scoring guide from the standards and the assignment** – Using the standards and the assignment, develop a scoring guide for this problem by following these steps: 4 is the highest score. Write the features of an excellent answer to this problem; 3 is the next highest score. Write the features of an answer clearly based on understanding of the concept with perhaps some minor errors that could be simple mistakes or typographical errors. Understanding of the problem and ability to apply it are obvious. A solid job, but not brilliant.
4. **We score the student work, using the guide** – Score the student work alone, first, using the scoring guide you've worked out together. When everyone has a set of scores, share them and reconcile them so that each team member roughly agrees. If you can't get complete agreement, at least decide between the papers that get a 4 or 3, and those that get a 2 or 1.
5. **We ask: Will this work meet the standards? If not, what are we going to do about it?** – THIS STEP AND THE FOLLOWING STEPS ARE THE MOST IMPORTANT IN THE PROCESS. People tend to think that they're done when they've got the work scored, but in fact all that was just preparation for answering the most important questions. Looking at the student work, please answer the following questions as a team: What does the student work tell us about learning in this classroom in this school? What do student's know and what are they able to do? Was the assignment well designed to help students achieve the standards?
6. **Implications for change: What are we going to do about it?** - The team should now answer this generic question: What should happen at the classroom, school, district, state levels to ensure that all students could achieve a score of 4 or 3 on assignments clearly aligned with the standards? The following are examples of actions that might be taken to improve learning: examining assignments to make sure that they are all aligned with standards, reorganizing curriculum and instruction, buying calculators and computers, etc. [Note: Original protocol contains more suggestions].

Figure 3. "Standards in Practice" protocol for reflecting on student work in collaborative professional development settings (Mitchell, 1997).

CRESST's Classroom Assignment Rubric and Professional Development

CRESST's classroom assignment rubric and method for collecting assignments could potentially serve both as a heuristic for designing high-quality assignments aligned with standards, as a structure to help guide discussion of assignments and student work. While research has not yet been conducted that focuses on the use and effectiveness of this method as a tool for teachers' professional development, as described before, results of CRESST's research indicate that students produced higher quality work when they were exposed to higher quality assignments. And at the secondary level at least, students scored higher on standardized tests of achievement when they had teachers who created more cognitively rigorous assignments that had clearer grading criteria. This suggests that

reflecting on the specific aspects of assignment quality in professional development could improve assignment quality (level of cognitive challenge and assessment criteria). Students who are exposed to more demanding work and higher quality grading criteria that provide them with better information about what they need to do to be successful on specific tasks, in turn, could potentially achieve at higher levels (see Figure 4).

To aid in the reflection process, teachers alone, with a coach, or in collaborative professional development settings could score assignments using CRESST-developed scoring manuals as a starting point. These manuals provide an example of a writing assignment and a reading comprehension assignment for each scale-point validated to date at Grades 3, 7, and 10 (see enclosed scoring manuals).

Teachers also could address questions pertaining to what they wanted students to learn as a result of completing a specific assignment task, how focused their goals were on the specifics of student learning, and how aligned their learning goals were with cognitively rigorous standards. They also could reflect on how challenging assignments were (e.g., if the assignment was to compare and contrast characters across stories, did students focus on surface level features only, or did they also engage with deeper content?). Additionally, teachers could reflect on whether an assignment was based on interesting and grade-appropriate academic content material (e.g., literature suggested in the California frameworks) and whether students were supported to write extended responses that utilized appropriate writing strategies (e.g., multi-paragraph essays for students at Grade 4 and above, etc.).

Teachers also could reflect on their grading criteria for assessing student work and discuss how clear the criteria were, how much information the criteria provided to students, and whether the criteria were aligned with their learning goals and standards. Teachers also could discuss ways to share their grading criteria with

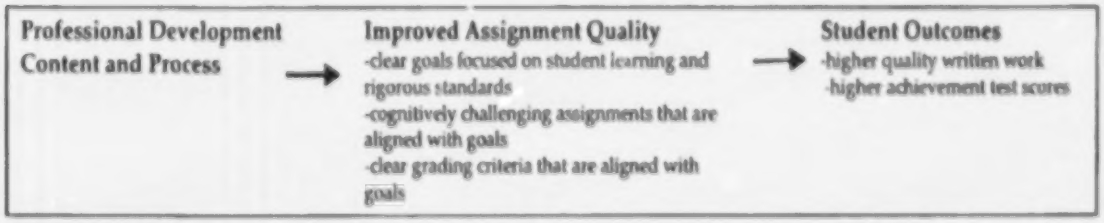
Professional Development
Content and Process
Improved Assignment Quality
-clear goals focused on student learning and rigorous standards
-cognitively challenging assignments that are aligned with goals
-clear grading criteria that are aligned with goals
Student Outcomes
-higher quality written work
-higher achievement test scores

Figure 4. CRESST's measure of assignment quality in professional development settings as a way to improve student learning.

students. Finally, teachers could look at the written feedback provided to students on drafts of their written work to see how their comments aligned with their grading criteria (see Figure 5 for a description of the potential questions teachers could ask themselves and each other to reflect on assignment quality).

While this method has yet to be used in professional development settings, teachers have indicated interest in using this tool. For example, CRESST introduced the rubric for assessing assignment quality to Critical Friends Group coaches, most of whom indicated that they would be interested in using the rubric in their group. Also, as described earlier, the LAUSD shared the criteria used to assess assignments with its participating teachers, many of whom reported that the rubric could be useful for them for reflecting on their practice. This method also recently was introduced to districts working with SERVE at the University of North Carolina. More research is needed, however, that focuses on how this tool could be introduced into collaborative professional development settings for teachers and the possible influence of using the classroom assignment rubric on teachers' instructional practice and student learning.

In summary, most researchers agree that traditional models of professional development are ineffective for supporting change in classroom practice and have advocated for school-based, collaborative professional development for teachers as an alternative to one-shot workshops and limited day institutes. While large-scale studies are few, the research indicates that these types of collaborative settings may positively influence teachers' instructional practices and student achievement. At the same time, however, it appears that existing protocols mostly focus on group process as opposed to classroom practice. Also, not all schools have access to instructional experts who are willing (and skilled enough) to serve as coaches for other teachers and who can provide more specific feedback to teachers on their lessons and assignments. It appears, therefore, that there may be a need for tools that help teachers reflect on the specifics of their practice in these settings (in addition to tools that support group process), and that the CRESST rubric may be useful in this capacity. Additionally, CRESST's method for looking at assignments and student work has the potential to help teachers align their everyday classroom practice with standards, and provide students with more challenging learning environments and higher quality feedback on their progress toward learning goals.

1. Instructional Goals Focused on Student Learning:

- What do I want students to learn as a result of completing the assignment? Are my goals focused on higher-order thinking skills? Are my goals aligned with standards and if so, which standards?
- Are my goals for this assignment focused on specific concepts I want students to learn, or are they mostly focused on student participation in activities?
- Do my assignment directions fully communicate my expectations to students? What could I add that might provide more guidance to students on what to include in their written work?

2. Cognitive Challenge:

- Does the assignment task require students to use higher order thinking skills (e.g., compare and contrast, identify themes, make predictions, solve problems, etc.)?
- Are students engaging with features of the text that go beyond surface-level details?
- Are students engaging with academic content material (e.g., selections from *Recommended Readings in Literature, Kindergarten Through Grade Twelve*)?
- Are students producing extended responses that are aligned with standards for writing strategies (e.g., multi-paragraph essays for students who are at fourth-grade and above)?

3. Grading Criteria:

- Are my assessment criteria informative to students with regard to what they would need to do to successfully complete the assignment?
- Did I share these criteria with students before they completed (or even started) the assignment?
- Did students participate in creating these criteria? How could I involve students more in creating assessment criteria?

4. Alignment of Goals and Task:

- Does the assignment task actually further my learning goals? Is there another way to design my assignment so that it better supports students learning of these concepts?
- Is the assignment aligned with standards? If so, which standards?

5. Alignment of Goals and Grading Criteria:

- Are students being assessed on the concepts I want them to learn as a result of completing the assignment (e.g., if the goal is for students to support their ideas with evidence, did my assessment criteria explicitly address the amount and quality of the evidence students used to support their ideas)?
- Is the feedback I give to students on rough and final drafts of their written work aligned with my learning goals (e.g., did my feedback to students focus primarily on mechanics, or did I also give them feedback that focused on content?

Figure 5. Sample questions for developing and reflecting on classroom assignments.

Conclusions and Directions for Future Research and Development

The major focus of CRESST's research and development efforts over the past four years has been on investigating the technical quality of the assignment ratings. Results so far indicate that rater reliability has been acceptable across study years, though more work may be necessary to develop scoring protocols for novice raters if this method is to be used in large-scale settings (e.g., in a large evaluation design or for accountability purposes). Results also indicate that the quality of classroom assignment ratings is significantly associated with the quality of observed instruction across most of the study years and across a range of learning

environments, providing support for the validity of the assignment ratings. Future research is needed, however, that focuses on the validity of these ratings in larger scale studies. Additionally, more research is needed that investigates whether the assignment ratings serve as valid and reliable indicators of classroom practice when teachers submit commercially produced assignments.

It appears likely that it would be necessary to collect as many as three or four assignments from teachers to obtain a stable estimate of quality. While teachers (who participated in the LAUSD study) were generally positive about the assignment rubric, they also reported that they needed a significant amount of time to produce assignments and collect student work. In the CRESST data collection efforts, teachers appeared to be the most satisfied when they received the assignment materials at the beginning of the school year (fall) and submitted assignments at the end of February and in early March. This timeframe appeared to give teachers the time they needed to create and implement assignments and did not compete with any major holidays. Also, it came well before the Stanford 9 testing in April.

Results also indicated that when students were exposed to high-quality assignments they produced higher quality work and received higher scores on standardized tests of achievement. The quality of assignments submitted at each of the study years, however, mostly tended to be of basic quality, and this was true even when teachers reportedly based their assignments on content standards. These findings raise important issues regarding teachers' interpretation and implementation of standards.

These results also suggest that improving the quality of teachers' assignments could possibly improve the quality of students' learning environments. The CRESST assignment rubric shows promise in that regard, though more work is needed to make the method even more "teacher friendly" than it is already, and to create guidelines for how to implement the method in collaborative professional development settings. CRESST's method also shows promise for helping teachers align their assignments more tightly with California's standards. It remains to be seen whether the rubric, protocol questions, and scoring guides alone would be enough to improve the quality of teachers' assignments. More importantly,

however, the question remains as to whether improving the quality of teachers' assignments alone would increase student achievement. These issues should be investigated in future research efforts.

References

Applebee, A. N. (1984). *Contexts for learning to write: Studies of secondary school instruction*. Norwood, NJ: Ablex.

Aschbacher, P. R. (1999). *Developing indicators of classroom practice to monitor and support school reform* (CSE Tech. Rep. No. 513). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Beck, I., & McKeown, M. (in press). Text talk: Capturing the benefits of read-aloud experiences for young children. *The Reading Teacher*.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*, 139-148.

Blume, J. (1972). *Tales of a fourth grade nothing*. New York: Dutton.

Briars, D. J., & Resnick, L. B. (2000). *Standards, Assessments—and What Else? The Essential Elements of Standards-Based School Improvement*. (CSE Tech. Rep. No. 528). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Brophy, J. (1992, April). Probing the subtleties of subject-matter teaching. *Educational Leadership, 49*(7), 4-8.

California Department of Education. (1998). *California Standards for the Teaching Profession*. Developed by the California Commission on Teaching Credentialing and the California Department of Education. Sacramento, CA: California Department of Education.

California Department of Education. (1999). *Reading/language arts framework for California public schools, kindergarten through grade twelve*. Developed by the California Development and Supplemental Materials Commission. Sacramento, CA: California Department of Education.

Cantrell, S., Lyon, N., Valdés, R., White, J., Recio, J, & Matsumura, L. (2001, June). *Pilot study report: The Local District Performance Measures*. LAUSD: LAUSD Program Evaluation and Research Branch.

Clare, L. (2000). *Using teachers' assignments as an indicator of classroom practice* (CSE Tech. Rep. No. 532). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Clare, L., & Aschbacher, P. R. (2001). Exploring the technical quality of using assignments and student work as indicators of classroom practice. *Educational Assessment, 7*, 39-59.

BEST COPY AVAILABLE

Clare, L., Valdés, R., Pascal, J., & Steinberg, J. R. (2001). *Teachers' assignments as indicators of instructional quality in elementary schools* (CSE Tech. Rep. No. 545). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Clark, C. M., & Yinger, R. J. (1977). Research on teacher thinking. *Curriculum Inquiry, 7*(4), 279-304.

Cohen, D. K., & Ball D. L. (1994). Relations between policy and practice: A commentary. *Educational Evaluation and Policy Analysis, 12*, 249-256.

Cohen, D. K., McLaughlin, M. W., & Talbert, J. E. (Eds.). (1993). *Teaching for understanding: Challenges for policy and practice.* San Francisco: Jossey-Bass.

Coalition of Essential Schools. (1996). Looking collaboratively at student work: An essential toolkit. *Horace, 13*(2). Retrieved March 11, 2002, from http://www.essentialschools.org/pubs/horace/13/v13n02.html

Cuban, L. (1990). Reforming again, again, and again. *Educational Researcher, 19*, 3-13.

Danielson, C. (1996.) *Enhancing professional practice: A framework for teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L. (2000, January). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives, 8*(1). Retrieved February 28, 2001, from http://olam.ed.asu.edu/epaa/v8n1

Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 76*, 597-604.

Duffy, G. (1981). Teacher effectiveness research: Implications for the reading profession. In M. Kamil (Ed.), *Directions in reading: Research and instruction. 30th yearbook of the National Reading Conference* (pp. 113-136). Washington, DC: National Reading Conference.

Estes, E. (1974). *The hundred dresses.* New York: Harcourt Brace Jovanovich.

Fuhrman, S. H. (Ed.). (1993). *Designing coherent education policy: Improving the system.* San Francisco: Jossey-Bass.

Fullan, M. (2000). The three stories of education reform. *Phi Delta Kappan, 81*(8), 581-584.

Gallimore, R., & Goldenberg, C. N. (1992). Tracking the developmental path of teachers and learners: A Vygotskian perspective. In F. K. Oser, A. Dick, & J.-L. Patry (Eds.), *Effective and responsible teaching: The new synthesis* (pp. 203-221). San Francisco: Jossey-Bass.

Gentile, C. (1992). *Exploring new methods for collecting students' school-based writing: NAEP's 1990 portfolio study.* Princeton, NJ: NAEP.

Goldenberg, C. (1992-1993). Instructional conversations: Promoting comprehension through discussion. *The Reading Teacher, 46,* 316-325.

Goodrich, H. (1996). Understanding rubrics. *Educational Leadership, 54*(4), 14-17.

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28,* 524-549.

Hiebert, J., & Stigler, J. (2000). A proposal for improving classroom teaching: Some lessons from the TIMSS video study. *The Elementary School Journal, 101*(1), 3-20.

Higuchi, C. (1996). *Improving student learning: High standards, standards-based curriculum, and standards-based assessment models.* Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Lieberman, A. (1994). Teacher development: Commitment and challenge. In P. Grimmett & J. Neufeld (Eds.), *Teacher development and the struggle for authenticity* (pp. 15-30). New York: Teachers College Press.

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis, 21,* 29-45.

Matsumura, L. C., Garnier, H., Pascal, J., & Valdés, R. (2002, April). *Classroom assignments as indicators of instructional quality.* Paper presented at the Annual Conference for the American Educational Research Association, New Orleans, LA.

Matsumura, L. C., Patthey-Chavez, G. G., Valdés, R., & Garnier, H. (in press). Teacher feedback, writing assignment quality, and third-grade students' revision in lower- and higher-achieving urban schools. *Elementary School Journal.*

Matsumura, L.C., & Steinberg, J. (under review). *Collaborative, school-based professional development settings for teachers: Implementation and links to improving the quality of classroom practice and student learning* (CSE Tech. Rep. No. 568). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

McDonald, J. (2001). Students' work and teachers' learning. In A. Lieberman & L. Miller (Eds.), *Teachers caught in the action: Professional development that matters* (pp. 209-235). New York: Teachers College Press.

McLaughlin, M. W., & Marsh, D. D. (1990). Staff development and school change. In A. Lieberman (Ed.), *Schools as collaborative cultures: Creating the future now* (pp. 213-232). New York: The Falmer Press.

McLaughlin, M. W., & Zarrow, J. (2001). Teachers engaged in evidence-based reform: Trajectories of teachers' inquiry, analysis, and action. In A. Lieberman & L. Miller (Eds.), *Teachers caught in the action: Professional development that matters* (pp. 79-101). New York: Teachers College.

Milne, A. A. (1927). *The house at pooh corner.* New York: Dutton.

Mitchell, R. (1997). Standards in practice. *The Education Trust* (Included in *Looking at Students work: A project of the Annenberg Institute for School Reform*). Retrieved February 12, 2002, from http://www.lasw.org/SIP_steps.html

National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for America's future.* New York: Author.

Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence?* Chicago: Consortium on Chicago School Research.

Newmann, F. M., Lopez, G., & Bryk, A. S. (1998). *The quality of intellectual work in Chicago schools: A baseline report.* Chicago: Consortium on Chicago School Research.

Newmann, F. M., Marks, H. M., & Gamoran, A. G. (1996, August). Authentic pedagogy and student performance. *American Journal of Education, 104*(4), 280-312.

Newmann, F. M., & Wehlage, G. G. (1993). Five standards of authentic instruction. *Educational Leadership, 50*(7), 8-12.

Onosko, J. J. (1992, April). Exploring the thinking of thoughtful teachers. *Educational Leadership, 49*(7), 40-43.

Palincsar, A. S. (1986). The role of dialogue in providing scaffolded instruction. *Educational Psychologist, 26,* 73-98.

Patthey-Chavez, G. G., & Clare, L. (1996). Task, talk, and text: The influence of instructional conversation on transitional bilingual writers. *Written Communication, 13,* 515-563.

Perkins, D., & Blythe, T. (1994). Putting understanding up front. *Educational Leadership, 51*(5), 4-7.

Peterson, A. (2001). NAEP/NWP study shows link between assignments, better student writing. *The Voice: A Newsletter of the National Writing Project, 6*(2), 1, 16-17.

Polacco, P. (1992). *Chicken Sunday.* New York: Scholastic.

Porter, A. C., & Brophy, J. (1988). Synthesis of research on good teaching: Insights from the work of the Institute for Research on Teaching. *Educational Leadership, 45*(8), 74-85.

Powell, A., Goldenberg, C., & Cano, L. (1995, April). *Assisting change: Some settings for teacher development work better than others.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Prawat, R. S. (1992, April). From individual differences to learning communities—our changing focus. *Educational Leadership, 49*(7), 9-13.

Project Zero. (2000). *Rubrics and self-assessment project.* Retrieved December 11, 2001, from http://pzweb.harvard.edu/Research/StuSA.htm

Rademacher, J. A., Cowart, M., Sparks, J., & Chism, V. (1997). Planning high-quality assignments with diverse learners. *Preventing School Failure, 42*(1), 12-18.

Resnick, L. B. (1995). From aptitude to effort: A new foundation for our schools. *Daedalus, 124*(4), 55-62.

Resnick, L. B. (1999, June 16). Making America smarter. *Education Week,* pp. 38-40.

Rothman, R., Slattery, J. B., Vranek, J. L., & Resnick, L. B. (2002). *Benchmarking and alignment of standards and testing.* (CSE Tech. Rep. No. 566). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Saunders, W., & Goldenberg, C. (1999). *The effects of instructional conversations and literature logs on the story comprehension and thematic understanding of English proficient and limited English proficient students* (Research Rep. No. 6). Santa Cruz, CA: Center for Research on Education, Diversity & Excellence.

Saunders, W., Goldenberg, C., & Hamann, J. (1992). Instructional conversations beget instructional conversations. *Teaching and Teacher Education, 8,* 199-218.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer.* Newbury Park, CA: Sage.

Slavin, R., & Madden, N. (1989). What works for students at risk: A research synthesis. *Educational Leadership, 46,* 4-13.

Spillane, J. P., & Zeuli, J. S. (1999). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. *Educational Evaluation and Policy Analysis, 21*(1), 1-27.

Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education.* New York: Summit Books.

Stokes, L. (2001). Lessons from an inquiring school: Forms of inquiry and conditions for teacher learning. In A. Lieberman & L. Miller (Eds.), *Teachers caught in the action: Professional development that matters* (pp. 141-158). New York: Teachers College Press.

Tharp, R. G. (1982). The effective instruction of comprehension: Results and description of the Kamehameha Early Education Project. *Reading Research Quarterly, 17*(4), 503-527.

Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context.* New York: Cambridge University Press.

Tyack, D., & Cuban, L. (1995). *Tinkering towards utopia: A century of school reform.* Cambridge, MA: Harvard University Press.

Tyack, D., & Tobin, W. (1994). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal, 31*(3), 453-479.

Vygotsky, L. S. (1978). *Mind in society. The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Appendix:

Teacher Assignment Cover Sheets

1 WRITING ASSIGNMENT COVER SHEET

1. Reading Material Information

Please write the title, author, and reading level of any reading material students read as part of this assignment.

Text Title	Author	Reading Level
a.		
b.		
c.		

2. Assignment Description

Describe the assignment in detail. Additionally, if applicable, please attach a copy of the assignment directions you distributed to students.

3. Learning Goals for Students

What were your *learning* goals for this assignment? Please describe the skills, concepts and/or facts you wanted students to learn as a result of completing this assignment.

4. Instructional Context

4a. How did this assignment fit in with your unit, or what you are teaching in your language arts class this month or this year?

OVER ⇨

4b. How long did students take to complete this assignment? _____

4c Approximately how many assignments do you give like this a year? [ ]

5. Grading Criteria

5a. Please describe your criteria for grading student work. If you used a rubric, please attach a copy of the rubric you used to grade student work for this assignment.

_____

_____

_____

_____

_____

_____

_____

5b. If you used a rubric to grade student work for this assignment, where did this rubric originate? Please check one or more of the following.

[ ] Self
[ ] Students
[ ] Teachers at my school
[ ] District, cluster or School Family
[ ] Published instructional program or teacher's guide
[ ] Other (please describe) _____

5c. Approximately what percentage of the students in your class performed at the following levels for this assignment?

_____% = Good to Excellent _____% = Adequate _____% = Not Yet Adequate

5d. What criteria did you use to decide what was "Medium" student work and what was "High" student work for this assignment? Please give specific examples from the papers you attach.

_____

_____

_____

_____

_____

_____

_____

2 READING ASSIGNMENT COVER SHEET

1. Reading Material Information

Please write the title, author, and reading level of any reading material students read as part of this assignment.

Text Title	Author	Reading Level
a.		
b.		
c.		

2. Assignment Description

Describe the assignment in detail. Additionally, if applicable, please attach a copy of the assignment directions you distributed to students.

3. Learning Goals for Students

What were your *learning* goals for this assignment? Please describe the skills, concepts and/or facts you wanted students to learn as a result of completing this assignment.

4. Instructional Context

4a. How did this assignment fit in with your unit, or what you are teaching in your language arts class this month or this year?

OVER ⇨

4b. How long did students take to complete this assignment? _____

4c Approximately how many assignments do you give like this a year? [ ]

5. Grading Criteria

5a. Please describe your criteria for grading student work. If you used a rubric, please attach a copy of the rubric you used to grade student work for this assignment.

_____

_____

_____

_____

_____

_____

_____

5b. If you used a rubric to grade student work for this assignment, where did this rubric originate? Please check one or more of the following.

[ ] Self
[ ] Students
[ ] Teachers at my school
[ ] District, cluster or School Family
[ ] Published instructional program or teacher's guide
[ ] Other (please describe) _____

5c. Approximately what percentage of the students in your class performed at the following levels for this assignment?

_____% = Good to Excellent _____% = Adequate _____% = Not Yet Adequate

5d. What criteria did you use to decide what was "Medium" student work and what was "High" student work for this assignment? Please give specific examples from the papers you attach.

_____

_____

_____

_____

_____

_____

_____

3 READING ASSIGNMENT COVER SHEET

1. Reading Material Information

Please write the title, author, and reading level of any reading material students read as part of this assignment.

Text Title	Author	Reading Level
a.		
b.		
c.		

2. Assignment Description

Describe the assignment in detail. Additionally, if applicable, please attach a copy of the assignment directions you distributed to students.

3. Learning Goals for Students

What were your *learning* goals for this assignment? Please describe the skills, concepts and/or facts you wanted students to learn as a result of completing this assignment.

4. Instructional Context

4a. How did this assignment fit in with your unit, or what you are teaching in your language arts class this month or this year?

OVER ⇨

4b. How long did students take to complete this assignment? _____

4c Approximately how many assignments do you give like this a year? ☐

5. Grading Criteria

5a. Please describe your criteria for grading student work. If you used a rubric, please attach a copy of the rubric you used to grade student work for this assignment.

_____

_____

_____

_____

_____

_____

_____

5b. If you used a rubric to grade student work for this assignment, where did this rubric originate? Please check one or more of the following.

[ ] Self
[ ] Students
[ ] Teachers at my school
[ ] District, cluster or School Family
[ ] Published instructional program or teacher's guide
[ ] Other (please describe) _____

5c. Approximately what percentage of the students in your class performed at the following levels for this assignment?

_____% = Good to Excellent _____% = Adequate _____% = Not Yet Adequate

5d. What criteria did you use to decide what was "Medium" student work and what was "High" student work for this assignment? Please give specific examples from the papers you attach.

_____

_____

_____

_____

_____

_____

_____

BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[X] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").